

Good Practices in Evaluating Tertiary PVE Programs: InFoEx Workshop, Berlin, September 19-20, 2019

Koller, Sofia

Veröffentlichungsversion / Published Version
Kurzbericht / abridged report

Empfohlene Zitierung / Suggested Citation:

Koller, S. (2020). *Good Practices in Evaluating Tertiary PVE Programs: InFoEx Workshop, Berlin, September 19-20, 2019*. (DGAP Report, 8). Berlin: Forschungsinstitut der Deutschen Gesellschaft für Auswärtige Politik e.V.. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67097-5>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:
<https://creativecommons.org/licenses/by-nc-nd/4.0>

by Sofia Koller
with contributions by
Prof. Adrian Cherney,
Dr. Björn Milbradt,
Dr. Tina Wilchen Christensen,
Amy-Jane Gielen,
and Motje Seidler.

BASE LINE ASSESSMENT

THEORY OF CHANGE

EVALUATION

OUTPUT

CLIENT

FUNDING

INPUT

PRACTITIONER

PRACTITIONER

FORMATIVE

IMPACT

CHALLENGES

GO

EVALUATION

OUTPUT

CHANGE

CASE NOTE

BASE LINE ASSESSMENT

ASSESSMENT

OUTPUT

MEASUREMENT

CHALLENGES

CLIENT

THEORY

OUTCOMES

INPUT

BASE LINE ASSESSMENT

OF CHANGE

CHANGE

CASE NOTE

FUNDING

GOOD PRACTICES

CASE NOTE

PRACTITIONER

CHALLENGES

THEORY OF CHANGE

MONITORING

OUTPUT

EVALUATION

OUTCOME

CHALLENGES

OUTPUT

PRACTITIONER

GOOD

GOOD

ABOUT THE PROJECT INTERNATIONAL FORUM FOR EXPERT EXCHANGE ON COUNTERING ISLAMIST EXTREMISM (INFOEX)

InFoEx is a joint project of the Migration, Integration, and Asylum Research Centre of the Federal Office for Migration and Refugees (BAMF) and the German Council on Foreign Relations (DGAP). Over the course of 2019 – 2020, InFoEx is collecting inspiring practices from practitioners working in tertiary prevention in Germany and abroad, as well as insights from academics conducting research in this field.

It is the project's objective to identify and generate empirical findings on processes of (de) radicalization, with a focus on their practical applicability for deradicalization efforts. To this end, the BAMF Research Centre initiated a consortium of research fellows who are embedded at local advice centers that work together with the BAMF Advice Centre on Radicalisation and various research institutions partnering with the BAMF Research Centre. These research fellows, along with the counselors working at the local advice centers, constitute the core stakeholders of InFoEx.

ABOUT THE WORKSHOP IN BERLIN, SEPTEMBER 19-20, 2019

Among the 30 participants were network partners of the BAMF Advice Centre on Radicalisation from civil society and government institutions, as well as practitioners and academics from Australia, Denmark, Germany, the Netherlands, and the United States. To align the workshop with the needs of its stakeholders, research fellows embedded at local advice centers in Germany shared – in agreement with practitioners at their local advice centers – specific information needs and questions regarding counseling work in tertiary prevention prior to the workshop.

CONTACT

Sofia Koller, Project Leader InFoEx, Email: koller@dgap.org

Content

Executive Summary	4
Key Recommendations	4
Introduction	5
1. Examples of Evaluations of Tertiary PVE Programs	7
2. What Does a Sound Evaluation Need?	10
3. Preparing for an Evaluation	11
4. Conducting an Evaluation	14
5. Outcomes of an Evaluation	16
6. Learning from Development Aid	17
Bibliography	18
Additional reading	19

Executive Summary

The transnational dimension of violent extremism calls for a comprehensive approach to prevention. International exchange of knowledge is crucial to enable an effective response. In addition, while more and more countries have increasingly focused on the prevention of violent extremism (PVE) (as opposed to purely repressive counter terrorism measures), actors need to be able to understand and demonstrate which measures work as well as how and why they do.

As part of the International Forum for Expert Exchange on Countering Islamist Extremism (InFoEx), an international workshop in September 2019 in Berlin addressed the issue of evaluating tertiary prevention of violent extremism, namely measures contributing to the disengagement, deradicalization, and rehabilitation of radicalized individuals. This Issue Paper presents the outcome of this exchange which was organized by the German Council on Foreign Relations (DGAP) in cooperation with the Federal Office for Migration and Refugees (BAMF). The workshop brought together some 30 participants from Australia, Denmark, Germany, the Netherlands, and the United States to discuss experiences, lessons learned, and inspiring practices on evaluation. After providing examples of how tertiary PVE programs in different countries have been evaluated, the paper presents challenges identified and good practices highlighted during international expert discussions. It also includes practical recommendations by international experts in the field.

From these findings, the following key recommendations result for those involved in the design, planning, funding, and implementation of evaluations of tertiary PVE projects and programs in Germany and abroad:

Key Recommendations

- 1** Establish a culture of evaluation among practitioners and funding institutions by including an evaluation perspective from the beginning, integrating it into the program's design, and providing opportunities to speak about what works and what does not.
 - 2** Consider using a participatory evaluation design and include multiple perspectives as well as ensure a climate of trust between practitioners and researchers so that they can communicate openly and at eye level.
 - 3** Provide long term funding and ensure sustainability for tertiary PVE programs and projects to help build relationships of trust with the concerned individuals. This will make it possible for evaluations to include a significantly delayed second round of interviews with clients and stakeholders.
 - 4** Provide opportunities for (further) professionalization of both practitioners and evaluators regarding challenges when evaluating tertiary PVE programs, projects, and measures.
 - 5** Ensure an intensified and structured exchange of good practices, relevant research, and policy responses between actors involved in evaluation, for example through national and international workshops and round tables.
-

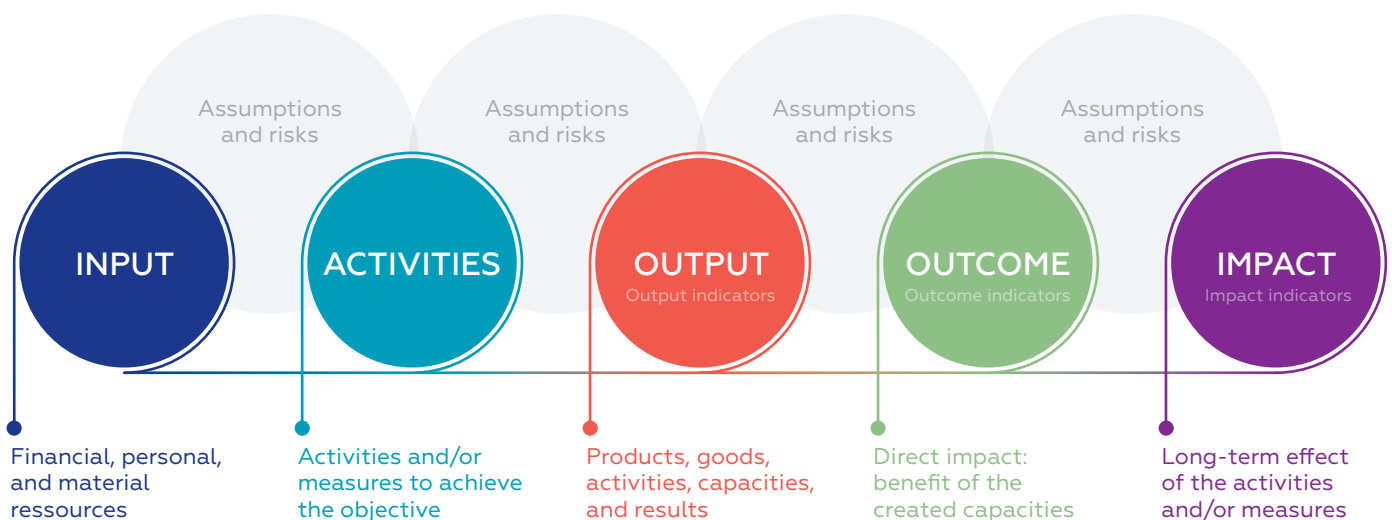
Introduction

Extremism and terrorism are not just national but international phenomena challenging the notion of a distinct internal or external security policy. In addition, the effectiveness of purely repressive counter terrorism measures has been called into question. Over the past years, European countries have increasingly invested in ‘softer’ efforts and measures to prevent and counter radicalization and (violent) extremism. Especially since the terrorist attacks in Paris in November 2015, it has become evident that the transnational dimension of violent extremism demands a comprehensive international and European approach to the prevention of violent extremism (PVE). Consequently, funding is being provided for a great variety of programs, projects, and measures. Yet to enable an effective response, structured and continuous knowledge exchange among actors in Germany as well as knowledge exchange between countries facing similar challenges is crucial.

The International Forum for Expert Exchange on Countering Islamist Extremism (InFoEx) provides a forum for this exchange on tertiary PVE : Stakeholders share a growing interest in measures that can contribute to the disengagement, deradicalization, and rehabilitation of radicalized individuals. So far, however, there is only limited understanding about whether, how, and why projects or measures work. Many factors can influence these processes, and it is not always possible to clearly identify the cause of changes and effects of a specific measure. Another debate evolves around the question of what constitutes success, how to measure it, and whether evaluation is limited to assessing the contribution towards a desired outcome.

There is a clear need for evaluating more tertiary PVE programs and sharing the results to establish “good practices.” This Issue Paper contributes to this objective by presenting the outcomes of expert discussions and debates documented during the third InFoEx workshop in September 2019. The

1. THEORY OF CHANGE



workshop's goal was to facilitate international knowledge exchange on evaluating programs and projects in the field of tertiary PVE¹. Around 30 participants from Australia, Denmark, Germany, the Netherlands, and the United States came together in Berlin to discuss experiences, lessons learned, and inspiring practices regarding evaluation. Integrating a needs-based approach², the workshop addressed topics such as the requirements for a sound evaluation; how to prepare and conduct an evaluation; what to do with the results of an evaluation; and how experiences in development aid can contribute to evaluating tertiary PVE work. The paper follows this structure: Each chapter summarizes highlights from the expert discussion documented during the workshop. Finally, international speakers have contributed their personal recommendations to practitioners working in tertiary prevention.

According to Dr. Björn Milbradt from the German Youth Institute (which evaluates large parts of the German program "Demokratie Leben!" [Live Democracy!]), evaluation should take a multi-perspectival view on deradicalization work. He observes that there is a great need in politics, research, and professional practice not only to find out whether something works, but specifically how that result is being achieved. Hence, a mix of methods is most useful for evaluating PVE programs, projects, and measures, since it is more flexible and can be tailored to specific objects of evaluation. Most evaluations are actually based on a mixture of different types, such as summative evaluations (focused on results), formative evaluations (using findings to modify an ongoing measure) or participatory evaluations.³ Impact evaluations focus on the effects or results of projects, namely outputs (observable achievements), outcomes (intended changes in the target group) as well as impacts (long-term individual and societal changes). This usually requires a theory of change that explains the functioning of a program or project and illustrates it in a simplified form (see figure 1).

Milbradt also mentions that at a technical level, evaluations can use qualitative methods such as interviews with project implementers, addressees or clients of tertiary PVE projects and measures, or quasi experimental designs (using a standardized questionnaire with addressees divided into an experimental group and a control group).

However, when reviewing existing evaluations on tertiary prevention, van der Heide and Schuurmann (2018) deplore that the "enduring scarcity of such assessments, particularly those based on first-hand information, remains a particular pressing issue" (p.197). According to Milbradt, this scarcity is due to the specific challenges of evaluating tertiary PVE programs: Here, experimental or quasi experimental evaluation designs reach their limits for technical and ethical reasons. For instance, the use of control groups is both difficult in practice and considered unethical by many practitioners working in the field. Milbradt considers it more practical to use qualitative sources and case monitoring, using indicators that should be as clearly defined as possible. In addition, (tertiary) prevention deals with a complex mix of ideological factors and highly individual case constellations. It is often carried out by multidisciplinary teams (including for example social workers and psychologists) which, absent long-term funding, lack sustainability. Trust between a counselor, a social worker, and a client is essential. It is the result of a frequently lengthy and highly sensitive relationship building process which raises issues of confidentiality and data protection. Involved stakeholders such as counselors at advice centers, funding agencies, and researchers tend to have very different positions and expectations of the employed methods and desirable outcomes.

Another issue is how to define and assess success in tertiary PVE. According to Milbradt, "ideological involvement as well as group memberships or contacts can also continue to exist; clients can always remain slightly below the level of delinquency in their actions or simply not get caught. Hence, a more complex view of success is needed in evaluation." For some researchers present at the workshop, long-term success means to restore a person's reflection capability by helping them think about what they are involved in. For others, success is not a question of statistical causality but of observing cognitive or behavioral change over time thanks to individualized treatment plans. Finally, there may be elements outside the program or project that influence behavior and need to be considered.

1 In the context of InFoEx, tertiary PVE programs are understood as all efforts aiming to disengage, deradicalize, decriminalize, and rehabilitate (violent) extremists in society and prison settings.

2 In preparation of the InFoEx workshop, research fellows working with local advice centers were asked to share – in agreement with practitioners – specific information needs and questions on the topic. This input was used to develop the workshop's format and content and to select relevant (international) speakers.

3 This paragraph is partly based on Dr. Milbradt's keynote speech at the workshop on September 19, 2019, in Berlin.

1. EXAMPLES OF EVALUATIONS OF TERTIARY PVE PROGRAMS

These challenges and the lack of empirical and representative studies and evaluations have been acknowledged for years (El Said 2015, Gielen 2018, Horgan & Braddock 2010, van der Heide & Schuurman 2018). At the same time, there is a growing body of literature and evaluations which aims to close this gap and contribute to a better understanding of the processes of disengagement, deradicalization, and rehabilitation. The following evaluations of tertiary PVE programs serve as examples of these efforts.



In the Netherlands, van der Heide & Schuurmann (2018) evaluated a specialized reintegration initiative run by the team Terrorism, Extremism, and Radicalization (TER) within the Dutch Probation Service (Reclassering Nederland, NR) over a period of 27 months. It is – so far – arguably the largest study of its kind in terms of the number and length of the interviews. In total, 72 semi-structured interviews were conducted in three rounds. The central research question was the extent to which “team TER’s activities have contributed to a lower chance of recidivism among clients with an extremist or terrorist background” (p.197). To this end, researchers evaluated three aspects. First, they looked at the program theory, assessing the goals of the program, the measurement of results, and the assumptions regarding cognitive and operational logic. Secondly, they conducted a process evaluation for the day-to-day implementation of the team’s work. Finally, they carried out a qualified impact evaluation to assess the program’s effectiveness as perceived by staff members and participants.

Regarding the program theory, van der Heide & Schuurmann found that the assumptions on how to bring about disengagement and deradicalization were theoretically sound. They argued, however, that the efficacy of the interventions underpinning the program’s operational logic could not be assessed empirically. Positive elements of the process evaluation were the use of theologians as external experts, good working relationships within the team, supervision of clients by teams of two, and the availability of a psychologist to help staff members with job-related stress. Finally, the impact evaluation showed that among 189 supervised clients, there were eight cases of terrorism-related recidivism and three

cases of unrelated recidivism. However, the authors pointed out that this assessment was problematic since no records were kept after the end of supervision. Nor did the program define a specific target for the recidivism rate. In conclusion, the authors note that “the program is held back by a lack of access to data on client’s long-term recidivism rates that could provide an, albeit imperfect, baseline for judging program effectiveness” (p.225f).



In Sweden, Christensen (2015) evaluated EXIT, a Swedish exit program for right-wing extremists that uses mentoring schemes with former right-wing extremists as coaches, therapeutic dialogue, and other activities to help the program’s clients to develop alternative worldviews, self-understanding, and identity. Christensen investigated the cultural and social sources which can be used to support an individual’s disengagement from an extremist group (p.15). For her research, she conducted 21 interviews with 15 people, enabling her to identify conditions that have helped former extremists to disengage. These conditions include making clients aware of the world view they had developed and using dialogue to add “grey tones to a white and black world view” (p.286). Another important aspect is the fact that clients and their coaches (since they are former extremists) have shared experiences. The “shame of having been involved in neo-Nazism or associated groups also renders an open and non-judgmental approach to the potential and present clients crucial for a relation to emerge” (p.287). Once trust is established, coaches are in a position to identify each client’s difficulties and motivations (hobbies, wishes for the future) individually. Finally, “clients need to learn to master social situations, which requires social competencies and depends on going through a situated learning process” (p.288). Christensen concludes that disengagement “seems to be possible by involving motivated individuals in learning methods embedded in everyday practice”, providing individuals with new social tools and supporting the development of higher self-esteem (p.289).



In Australia, Cherney (2018a, 2018b) evaluated the proactive integrated support model (PRISM), a pilot intervention delivered by the Corrective Services New South Wales and focused on convicted terrorists and prison inmates identified as at risk of radicalization. Data is derived from interviews with program staff, corrective services personnel as well as clients of the intervention (i.e. convicted terrorist and radicalized inmates and parolees). A first, process evaluation found that offenders reported to be motivated to participate in PRISM to demonstrate that they are not radicalized (anymore), which in turn created an opportunity to begin a process of self-reflection with them. Participants benefitted by engaging in self-reflection on the reasons for their radicalization. They also received help to cope with their time in custody and prepare for release. Regarding ideological components, the intervention aimed to “promote a plural and more in-depth understanding of Islam” (2018a, p.16). A second, larger evaluation was based on 38 interviews as well as the case notes of 15 clients and contributed to an understanding of the intervention’s goals and the methods of achieving them: “While varied, the needs being addressed through the PRISM intervention help to improve psychological coping, promote self-reflection and offence insights, and focus on religious mentoring and the development of prosocial supports and activities” (p.21). Cherney furthermore drew attention to the importance of program staff to capture “hooks for change” and recognize their significance for supporting clients in their disengagement process. At the same time, some of the benefits that were reported by PRISM clients could be considered “relatively standard forms of assistance characteristic of many in-custody rehabilitation programs. This draws attention to the fact that some of the needs of radicalized/terrorist inmates are not all that different from those of ‘mainstream’ offenders, particularly when it comes to their reintegration” (2018b, p.21-2).



In Sri Lanka, Webber et al (2018) examined the effectiveness of the Sri Lankan rehabilitation program for former members of the terrorist organization Liberation Tigers of Tamil Eelam (LTTE). In a first study, they surveyed former LTTE

members, comparing the responses of 490 beneficiaries of the full rehabilitation program to 111 beneficiaries who received only minimal rehabilitation. Their questions touched on topics such as embeddedness in the organization, attitudes towards deradicalization, feelings of insignificance as well as extremism (measured in terms of attitudinal support for LTTE’s ideology and approval for the use of violence to achieve these goals). The first study demonstrated that “beneficiaries receiving full rehabilitation reported increasingly lower extremism across one year and showed greater reduction than those only receiving minimal treatment. (...) Feelings of insignificance acted as a mechanism underlying this reduction. The rehabilitation program thus appeared to successfully address the lack of significance in participants by providing them with alternative mechanisms for earning significance” (p.8). In a second study, the authors surveyed 179 former LTTE, and 144 Tamil community members. The results suggest that the changes achieved through rehabilitation were long-lasting, since LTTE members who were beneficiaries of rehabilitation were significantly less extreme than the matched community members. Furthermore, the findings drew attention to the “critical role of personal significance in effective rehabilitation. Positive attitudes toward rehabilitation and greater participation in rehabilitation programming buffered participants against feeling insignificant, even after they had graduated from rehabilitation and returned to their communities” (p.13).

Another example from outside the European context is the evaluation of the Serendi Rehabilitation in Mogadishu, Somalia, for ‘low-risk’ former members of Al-Shabaab. It provides insights into why they joined Al-Shabaab, how and why they disengaged, and what their experiences of reintegration into the community were (Khalil et al 2019).



Especially Germany is doing more work on evaluations. The evaluation of Uhlmann (2017) at the Research Centre of the Federal Office for Migration and Refugees (BAMF) focused on the work of the BAMF “Advice Centre on Radicalization” and its then four civil society partners and laid the foundation for a comprehensive impact analysis. In recent years, a growing number of German exit programs (for both right wing and Islamist extremism) have been evaluated (Möller et al 2015, Möller & Neuscheler 2018, Schuhmacher 2018). Several more are currently being evaluated or will soon be.



Beside a growing body of research and more tertiary PVE programs evaluated at a national level, for example in France and Germany (Köhler 2019), knowledge about relevant topics is also increasingly being shared internationally (RAN 2018). For example, RAND Europe conducted a five-year study on the design and conduct of evaluations of CT and PCVE policies in the Netherlands and abroad (Bellasio et al 2018). Furthermore, the German Federal Ministry of the Interior, Building, and Community is currently preparing “Practical Guideline for Policymakers and Practitioners for Planning, Implementing and Following Up on Evaluations in Exit / Intervention Measures”. This guideline is the result of an EU member state project led by Germany in cooperation with the European Commission. The final guidelines are expected to be published in early 2020.

After having published guidelines for the evaluation of P/CVE programs and interventions (RAN 2018) and a paper on evaluating disengagement, deradicalization, and resocialization efforts (RAN 2019b, available in English, French, and German), the Radicalization Awareness Network (RAN) has recently also published a manual for peer and self-review in exit work (RAN 2020). It includes practical guidelines (on preparing the review, the review process, and how to follow up on results), a roadmap for implementation, and a list of definitions.

2. WHAT DOES A SOUND EVALUATION NEED?

While these studies identify crucial elements, they also highlight the difficulties of evaluating interventions. At the workshop, experienced evaluators and researchers, employees from state agencies as well as practitioners working in the field of tertiary prevention in Germany and abroad discussed what they consider important for conducting a sound evaluation.

Highlights from the Expert Discussion

Researchers and practitioners at the workshop largely agreed that all parties involved in an evaluation should have **a common understanding about the evaluation's objectives** and agree on a theory of change. There should also be a clarity of roles, competences, and responsibility of all stakeholders involved in the evaluation.

Many researchers consider having a baseline assessment as a prerequisite for any program: Practitioners should develop **a baseline assessment**⁴ to be able to gauge the client's position at the beginning of an intervention. However, it was pointed out that disengagement and deradicalization are processes that are not necessarily linear. They can take several years and include ups and downs. Therefore, some social workers do not use interviews, but a structured reflection on the client. Some use a mind map to identify the client's characteristics, recognize what can be worked on based on behavioral psychology, and assess change. Others use mind maps based on checklists and risk factors. Some counselors considered it helpful to ask clients what they want to talk about, what is going well, and what can energize them. This could enable the client to cooperate, start a reflection process, and help identify challenges as well as motivations.

Some evaluators also insist on the need for practitioners to ensure **consistent data collection**. They see this as helpful not only for external evaluation purposes but also to internally monitor change over time, since the duration of a client's participation in a deradicalization program can vary from a few months to several years. At the same time, one perception was that practitioners may be reluctant to share data on their cases with evaluators, for example due to data protection concerns.

A monitoring component to enable formative evaluation was also considered important, since an evaluation could be used to improve a program, for instance by adapting a project's focus. Monitoring changes in the client attitude and participation could also be very useful: With such a component, adjustments to a clients' intervention plan can be made as he or she progresses through the intervention. This may be important as intervention goals in the beginning of a participation can differ from when clients (or offenders) approach being released or ending their participation in a program. For this purpose, monitoring should include basic case note processes and data collection.

Most experts agreed that since deradicalization is considered just as much of a process as radicalization, a **significantly delayed second round of interviews** with clients and stakeholders is important to determine the (contributing) impact of the intervention on mid- or long-term behavioral or attitudinal changes of the client.

Similarly, a **long-term relation** was deemed necessary to understand the client's current situation and to assess the program's impact. In an effective system, exit workers should have a good relationship with the other stakeholders and be able to gain the trust of the client and her or his family, so that several interviews can be conducted over a longer period. This approach requires long-term projects and long-term contracts for social workers. Failing that, it may not be possible (in terms of information and time) to follow up on clients' development.

Regarding **risk assessment**, opinions tended to differ, and many exit workers expressed skepticism toward using standardized risk assessment tools. However, some make use of such tools to complement their professional assessment of the possible risk a client presents. According to one position, a continuous and consistent risk assessment is in the counselors' interests: Not only does it support exit workers in their job, but it can also help demonstrate their work's consequence and impact. Others believe that there is no need for risk assessment as such since radicalization does not necessarily entail security risks, and most extremists do not engage in violence. Then again, counselors may routinely include risk assessments in their work without using that expression. Finally, while counselors often base their assessments on academic education as well as years of experience, these may still be difficult to standardize.

⁴ A baseline assessment is "an assessment of the current situation vis-à-vis the [project's] current framework and should be conducted before the [project or measure] starts. It contains the specifications against which the outputs are measured. (...) Without a proper baseline assessment, measuring (...) outcomes is difficult, as outcomes are described as a measurable change that is an indirect result of an intervention (RAN 2019a, p.6).

3. PREPARING FOR AN EVALUATION

Having considered key components for a sound evaluation, the following chapter focuses on proposals for preparing the evaluation of a tertiary PVE program or project.

Highlights from Expert Discussion

The many stakeholders in the field of tertiary prevention tend to have various, sometimes conflicting goals and objectives, which are reflected in their approach to evaluation. It is not only in Germany that exit work can be perceived as a political mine field: Funding institutions (such as state ministries) are seen to favor statistics to demonstrate the success of (tertiary) PVE programs; social workers appear to focus on the reintegration of an individual in society; and security institutions tend to define success as being able to reduce the security risk. There can also be a difference between theoretical and practical conditions.

Sometimes, it will take until the start of an evaluation to notice that not all the relevant goals of a program were formulated at the beginning (see chapter 2). Secondly, measuring the desired effect (or impact) is challenging, and some experts even considered it impossible to attribute any kind of change to the intervention with certainty. There also appeared to be a lack of common understanding between practitioners and evaluators concerning the metrics of change, for example for measuring cognitive change. Some argued that the emotional changes that can occur during an intervention should be researched and evaluated more from a client's perspective. One proposition was to measure the probability of impact and consequently assess whether the program has a maximum probability to generate impact. There was also a debate about the appropriate starting and ending point for measuring change. Others argued that at least the structural integrity of a program is quite easy to measure. This includes assessing whether a project is based on a theory of change and clearly defined objectives.

To respond to these challenges, the following proposals and good practices were shared:

- **Plan together:** Ideally, questions regarding evaluation, monitoring, and data collection should be considered right from the start when designing a new program or project.
- **Ensure a process of co-creation:** (External) Researchers, practitioners, and funding institutions should be included as partners in setting up the research design. This approach ensures that all stakeholders have the possibility to communicate their positions and objectives and participate in the evaluation process, which would help establishing the measures' improvement as a common goal. Researchers, for

example social scientists, should be included from the start to ensure that scientific expectations are communicated. They should be able to formulate indicators for output and outcome to allow for a proper and thorough evaluation. This would also mean that intervention providers know from the beginning what data they need to collect. For example, a questionnaire could be included in the intake procedure to help with the baseline assessment. Some considered it helpful to task an external research institution with the evaluation, since it can provide a different perspective, relevant expertise, and enough resources to carry out the evaluation. Finally, practitioners should be directly involved, since they have better knowledge of what works or doesn't work with an individual.

- **Develop a common understanding and agree on terminology:** It has proven useful for the researcher, the project lead, the civil society organization, and the funding institution to jointly discuss and determine the evaluation's design. It was recommended that funding institutions, implementors, practitioners, and, if possible, clients agree on a base line assessment, a theory of change, definition and metrics of success, objectives, methods, needs, perceptions, as well as the use of case notes. It is also considered helpful to be transparent about the meaning of terms for defining and refining core concepts. Developing a common vocabulary and understanding needs a constant process of reflection but can improve communication among stakeholders. For example, instead of using the controversial term 'deradicalization', other paradigms can be used, such as the 5 Cs of positive youth development (PYD): competence, confidence, connection, character, caring (discussed for example by Lerner et al (2005)).
- It is also considered important to manage expectations and fears and to be aware of needs, especially of funding institutions. The complexity of the field of tertiary PVE should be made transparent to society and political actors.

Long-term funding: Since tertiary PVE programs often involve one-on-one interventions, they tend to be more resource-intensive in terms of funding and time than primary or secondary PVE programs which can work with larger groups. Monitoring and evaluation add to that. Another issue mentioned was the potential for tensions between stakeholders during an evaluation process caused by the fact that the outcome could affect the program's funding: While funding institutions have to ensure and prove that public money is spent on effective measures, evaluations need to be conducted in an environment that allows programs to be adjusted and improved in accordance with the evaluation results. Some practitioners feared negative consequences and "being punished" if they share that an intervention did not yield the desired results.

Finally, exit workers voiced concerns about the general lack of long-term funding which leads to a rapid turnover of personnel since many positions are only funded for one year. On the one hand, this means constantly rebuilding the professional networks within and between institutions which also has an impact on trust building with evaluators. On the other hand, it becomes more difficult to assess a project's long-term impact, as such an evaluation would have to be based on data from organizations that have gone through many personnel changes since the projects at stake were finalized. At the same time, evaluations are needed to decide which actor gets long-term funding. The following proposals were made:

- **When preparing for an evaluation, long-term funding is important:** Evaluations should cover a period of at least two years to be able to adequately measure outcomes. Researchers at the workshop also considered it important to plan for enough time after the evaluation so that stakeholders can reflect together on the identified challenges and issues and the consequences for their daily work. The German Youth Institute (DJI) for example has annual meetings to reflect on their researchers' approach.
- **To establish trust, evaluators should communicate** and ensure from the beginning that it is possible for all the involved actors to talk honestly about their fears, uncertainties, or criticisms in relation to the evaluation as well as the intervention itself. In this regard – and since evaluations add to practitioners' considerable workload – it could be considered helpful for evaluators to communicate how the evaluation can support and improve the practitioners' work. Other proposals included to work toward a "positive wording" of evaluation.

One of the most controversial topics was the collection and sharing of data. On the one hand, international researchers agreed on the importance of collecting data from the very beginning of a program or project. They also mentioned that the lack of consistent structures for case notes poses a problem for evaluators when comparing cases. On the other hand, social workers expressed skepticism towards recording (even anonymized) data, either about their clients or about their working processes. They fear this could lead to a distorted categorization of clients' personal and ongoing processes. For the same reason, they are often uncomfortable or even unwilling to share internal data. Another reason is that protection laws prohibit sharing personal – and often sensitive – data without the concerned person's consent.

The following proposals and experiences were shared:

- It can be helpful for the advice center, the civil society organization, and the funding government agency to discuss and establish meaningful and carefully defined parameters of documentation together. The objective would be to develop case documentation and documentation on practitioners' work that is supportive of both the practitioners' and the state agencies' work. This process can take time: In one case, it took three years to develop a documentation that was considered advantageous by both parties.
- **A consistent case note structure** that captures client participation, progress, and other measures of change was recommended. Ideally this should be designed before the start of the program. If this has not been done, experienced evaluators agreed that an important step is to identify whether case notes of practitioners are available, and how consistent they are.
- It was also mentioned that personal relationships between stakeholders were more important than the legal framework when it comes to data sharing.

RECOMMENDATIONS

By Amy-Jane Gielen, Consultant, A.G. Advies & PhD Candidate, University of Amsterdam, the Netherlands:

- Always include researchers prior to the implementation of a CVE program. This enables them to do a baseline assessment before implementation.
- Formulate indicators when designing the CVE program. What criteria should be met before we can speak of success?
- Use a risk-assessment model before, during, and after the tertiary CVE program. My preference is the Pro-Integration Model developed by Kate Barrelle (see p.15).

EXAMPLE REALISTIC EVALUATIONS

Apart from the general lack of evaluations of exit programs, there is also an expressed need to evaluate efforts to deradicalize, disengage, reintegrate, or rehabilitate women involved in violent extremism: Gielen (2018) proposes a model for conducting a realistic evaluation of the effectiveness of exit programs for jihadist women. The four steps include:

Developing a hypothesis on contextual conditions and mechanisms: Relevant contextual conditions can include for example physical space (such as an advice center or prison), geographical location (such as Amsterdam or Den Haag), target audience, individual capacities of key actors, interpersonal relationship between intervention providers and the client as well as broader institutional setting. Relevant mechanisms can include mentoring, religious and ideological support, practical support in establishing a stable environment and daily routine, psychological support and counselling, family support as well as administrative and legal measures.

Analyzing possible outcome patterns and related measurable indicators: Concerning an exit program for jihadist women who have attempted to travel to the so-called Islamic State territory, one indicator could be further attempts by the client to travel. Gielen also proposes to use Barrelle's (2015) Pro-Integration Model (see page 15).

Using a multi-method data collection: This can include for example face-to-face, questionnaire-based interviews; desk research involving for instance police and municipal registries or child protection services files; and interviews with stakeholders such as social workers or clients.

Developing a more refined theoretical model: This model should highlight how the contextual conditions and mechanisms of exit strategies "lead to specific outcome patterns" (p.464).

Developing this model should "help to answer the research question 'what works, for whom, how and in what circumstances in exit programs for female jihadists'" (p.456).

4. CONDUCTING AN EVALUATION

Highlights from the Expert Discussion

During the discussion, there was no agreement on the definition, role, and objective of monitoring when conducting an evaluation. While monitoring case work relates to evaluating the case work of single cases, monitoring the process of a program relates to monitoring standards, outcomes, etc. The latter was understood as a continuous process in which systems and instruments are used to assess progress, focusing on processes, activities, and outputs instead of outcome and impact. The results of this kind of monitoring can be used to adapt projects to increase the chance that the desired outcome will be achieved.

- In this context, a project aimed at improving cooperation between the community and the police was mentioned that included a built-in monitoring component as part of the evaluation. This way, interim results were used to inform police, and the project could be adapted where necessary. The monitoring part was intended and understood to be not just a control but also an information mechanism.

Data collecting and data sharing also remains a key challenge when conducting an evaluation. As van der Heide & Schuurmann (2018) have noted in their evaluation of an exit program run by the Dutch Probation Service, “systematically gathering quantifiable data on the risk of terrorism-related recidivism is likely to be the foremost challenge for reintegration programs” (p.226).

- When conducting an evaluation, researchers can be important facilitators between state agencies and practitioners to explain problems and objectives. Involving both external and internal researchers can be complementary for an evaluation, since an external researcher may be able to better identify blind spots. Some practitioners argued that in many cases, it is easier for internal researchers to build trust with clients and other actors in the field. Practitioners should also be closely involved in the ongoing process.

- Regarding communication, a common language would ensure that all actors understand questions and results.

RECOMMENDATIONS

By Dr. Tina Wilchen Christensen, Assistant Professor, Aarhus University, Denmark:

- Programs need to be based on a clearly defined theory of change in order to focus and evaluate the program.
- Programs should have a monitoring component to be able to investigate if the methods applied and the practices established within the programs provide the right support for the people in the program according to the theory of change. A monitoring component also helps employees develop a reflective approach to their own practices, which in turn makes it possible to adjust the theory of change in accordance with the outcome of the monitoring component.
- Programs should recruit people who have a thorough understanding of the field and the context in which they work. They should be sensitive toward the many paths that lead people to join violent extremist groups such as structural, individual, and coincidental matters, and understand that people hold different positions in violent extremist groups.

RESEARCH PERSPECTIVE: PRO-INTEGRATION MODEL BY KATE BARRELLE

To assess (the extent of) disengagement, Barrelle (2015) proposes a five-domain, three-level model of disengagement called the Pro-Integration Model. In each domain, there are three levels of social engagement which can be minimal, cautious, or positive (see figure 1). Barrelle concludes that “social relations are the vehicle through which most change occurs, so consideration of who a person spends time with is critical. Coping skills and self-care are necessary for an individual to move from surviving to thriving in society” (p.140): “In short, disengagement is actually about engagement somewhere else” (Barrelle p.133).

Barrelle’s model is seen as useful by several researchers. For example, Cherney & Belton (2019) have tested Barrelle’s PIM as an evaluation tool to examine outcomes for clients who participated in the Proactive Integrated Support Model (PRISM) intervention in Australia. The following example from Gielen (2018, p. 468) describes the case of a woman showing positive levels of social engagement in each domain:

THE PRO-INTEGRATION MODEL

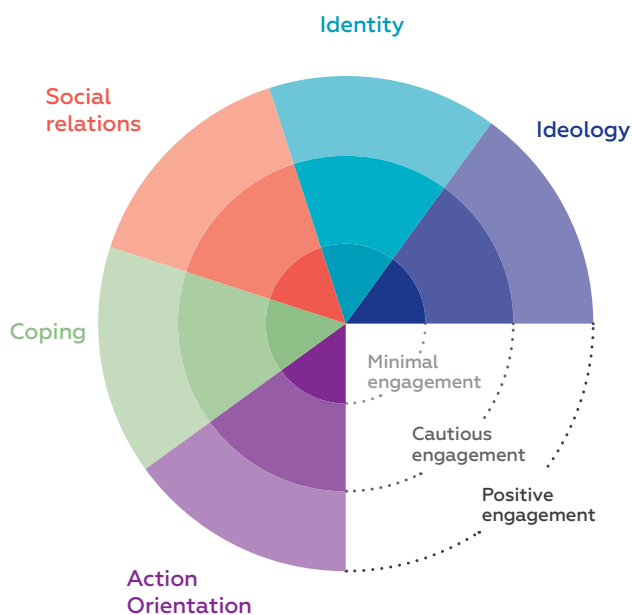


Figure 2: Adapted from Barrelle 2015 (p.135)

IDEOLOGY

Does not hold radical views, respects other (world)views, even gave her intervention provider a Christmas card, focuses on moderate law school and scholars.

IDENTITY

No identification with extremist group, proper sense of self and life history, no longer categorizes in ‘us’ and ‘them,’ and no longer uses extremist recruitment narratives for explaining why she wanted to join ISIS.

SOCIAL RELATIONS

Positive relationship with family, friendly relationship with non-Muslims, and no contact with former extremist network.

COPING

Able to address personal issues, understand push and pull factors and the events that triggered the desire to join ISIS, and undertake meaningful activities by going back to school and getting a part-time job.

ACTION ORIENTATION

Does not consider violence a legitimate method, no longer wants to travel to join ISIS, but focuses on her future in the Netherlands and positive action and participation by other means (e.g. joining the debate team at school).

5. OUTCOMES OF AN EVALUATION

After an evaluation has been carried out, an important aspect is what to do with the results and how to communicate them to different target groups. Especially advice centers voiced the fear that in a field where many actors compete for funding, negative evaluation results could lead to a negative backlash.

Highlights of Expert Discussion

Generally, since there have not been many evaluations of tertiary PVE programs carried out so far, there appears to be a lack of experience and knowledge regarding the outcomes of evaluation. A German researcher argued that evaluations are often conducted like short-term research with little attention given to the application of the outcome. Challenges and good practices regarding this topic can be divided into the following two aspects:

Communication of results to stakeholders

Tertiary PVE programs often involve multiple actors who contribute to different aspects of an intervention. Participants argued that these actors need to know if their contribution was leading to a change in the client's behavior or thinking, or if they should change their approach in any way. Evaluations could help inform such decisions and support changes to a program. One state agency representative said it was important to not overemphasize numbers and deficits. Instead, evaluation should be considered as a tool that can highlight the value and quality of exit work. However, practitioners voiced the concern that details of their working methods could be leaked to the public and/or clients and that they could lose their competitive advantage once the evaluation's results were published. At the same time, stakeholders such as funding institutions saw the benefit of communicating evaluation results since that could help promote longer-term political and financial support for a program. The following proposals were made:

- Proper communication of evaluation results can help secure the support of key communities and agencies. These actors can play an important role in supporting the reintegration of a formerly radicalized individual or inmate when released from prison. Public communication can also make exit programs more visible for people who wish to disengage from an extremist group.
- Several experts argued that after the end of the evaluation, the project should be adapted or further developed according to the evaluation results. After 6 or 12 months, the practitioners should take stock of those changes together with the evaluators.

- One proposal was that an internal report should be published first, including definitions, methods, data, and a short summary. After approval by all stakeholders, these results could then be communicated to the public to provide context and understanding.

- A good practice that was shared referred to a case where incorrect or politically problematic formulations of a report were discussed with the actors involved and changed prior to publication. Theoretically, the completed evaluation could have been published without an official release by the lead partner.

Communication of results to the public

Communicating an evaluation's results to the public can have an important impact on the reputation of a program and organization. Stakeholders shared their concern that an evaluation might show that the program was not working which would lead to public and political criticism. One researcher pointed out that there could be a reluctance to share information regarding terrorist convicts with the broader public. Also, the complete program could be branded as a failure if it became known that even one terrorist inmate, who was a client of a tertiary PVE program, re-engaged in extremism after being released. Yet experts also largely agreed that talking to the public is crucial to raise public awareness and to communicate a program's success as there is low tolerance of failure within the public sphere.

Participants argued that there seems to be a lack of public understanding about specific challenges of tertiary prevention, for example that no miracles can be expected, and that recidivism is possible. There was also a perception of low tolerance in the political sphere, since politicians are under pressure for justifying expenditure and providing security to the public. Instead of blaming the program, one recommendation was to establish a culture of talking about failure and sharing this with the public. Then again, it was mentioned that some political or societal groups might misuse an evaluation's outcome (success or failure) to support their political statements.

Since tertiary PVE programs are normally funded by the state and thus through taxpayers' money, one opinion was that the public needs to be informed of programs and their results.

- Evaluators and practitioners should work hand in hand in order to avoid distrust. Instead, they should promote trust in practitioners' work and expertise.

The question of how to define success and assess impact of tertiary PVE programs or projects remains challenging. It is crucial to be able to properly communicate this challenge.

RECOMMENDATIONS

By Prof. Adrian Cherney, Associate Professor, University of Queensland, Australia:

- Look at a variety of outcomes: Look at both behavioral and attitudinal change as well as at compliance with the intervention.
- Acknowledge that outcomes and success will vary with different clients: Success will not look the same for every individual, and indicators need to capture such variation.
- Examine both quantitative and qualitative outcomes: Quantitative and qualitative data on program outcomes should be collected relating to clients. Insights from staff can also be useful.

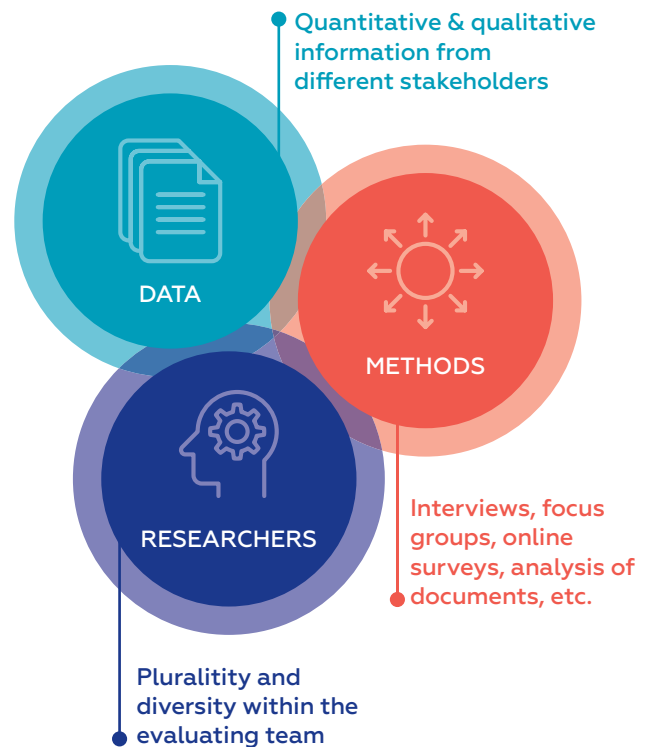
6. LEARNING FROM DEVELOPMENT AID

As evaluations of tertiary prevention are still rare, development aid and peace building projects can provide important insights and good practices regarding the reassessment of tools and decision-making processes. During the workshop, the benefits of developing a common understanding of the project's objectives and of clearly expressing assumptions on causes and effects as well as on the program's logic were mentioned.

In addition, it was said that the meaning of 'impact' often remains abstract. Concrete and tangible indicators should be chosen to arrive at an operational definition of the intended impact. Another good practice mentioned was triangulation: Data, methods, and research are used for a contribution analysis (see figure 3). This increases the variability in data gathering.

However, there are also limits to the use of such experiences. Some practitioners argued that ethical questions need to play a bigger role in practical tertiary PVE work and in the evaluations. This should include further reflection on how to deal with stigmatization. For instance, for an evaluation after release, it may not be appropriate for evaluators to question the employer of a client. It was proposed to not focus on causality but instead analyze the theory-led contribution towards the measure's objective.

TRIANGULATION



Own graphic, adapted from presentation by Motje Seidler, Syspons GmbH

RECOMMENDATIONS

By Motje Seidler, Consultant, Syspons GmbH, Germany:

- Develop a theory of change of the program to be evaluated, which includes inputs, activities, outputs, outcomes, and long-term impacts that the program seeks to achieve.
- Find a positive wording for the intended outputs, outcomes, and impacts of the program, so that changes can be measured more easily, for example "competencies of actors are strengthened" instead of "actors do not feel insecure anymore." Subsequently, operationalize the intended outputs, outcomes, and impacts.
- Apply several forms of triangulation in the evaluation process, such as data, methods, and researcher triangulation.

BIBLIOGRAPHY

- Barrelle, K.: "Pro-Integration: Disengagement from and life after extremism," *Behavioural Sciences of Terrorism and Political Aggression* (2015), 7(2), pp. 129-142.
- Bellasio, J., Hofman, J., Ward, A., Nederveen, F., Knack, A., Meranto, A.S. & Hoorens, S.: "Counterterrorism evaluation. Taking stock and looking ahead." Research report. RAND Europe (2018).
- Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung: Evaluierungskriterien für die deutsche bilaterale Entwicklungszusammenarbeit. Eine Orientierung für Evaluierungen des BMZ und der Durchführungsorganisationen, (2006).
- Cherney, A.: "Supporting disengagement and reintegration: qualitative outcomes from a custody-based counter radicalisation intervention," *Journal for Deradicalization* (2018b), Vol.17, pp. 1-27.
- Cherney, A.: "Evaluating interventions to disengage extremist offenders: a study of the proactive integrated support model (PRISM)," *Behavioral Sciences of Terrorism and Political Aggression* (2018a).
- Cherney, A. & Belton, E.: "Evaluating Case-Managed Approaches to Counter Radicalization and Violent Extremism". An Example of the Proactive Integrated Support Model (PRISM) Intervention, *Studies in Conflict and Terrorism* (2019).
- Christensen, T.W.: "A question of participation – Disengagement from the Extremist Right. A case study from Sweden," Thesis (2015), Roskilde University.
- El Said, H.: "New Approaches to Countering Terrorism. Designing and Evaluating Counter Radicalization and De-Radicalization Programs," Hampshire, United Kingdom: Palgrave MacMillan (2015).
- Gielen, A.: "Exit Programs for Female Jihadists. A Proposal for Conducting a Realistic Evaluation of the Dutch Approach," *International Sociology* (2018), pp. 454-472.
- Horgan, J. & Braddock, K.: "Rehabilitating the terrorists? Challenges in assessing the effectiveness of de-radicalization programs," *Terrorism and Political Violence* (2010), 22(2), pp. 267-291.
- Khalil, J., Brown, R., Chant, C., Olowo, P., & Wood, N.: "Deradicalisation and Disengagement in Somalia. Evidence from a Rehabilitation Programme for Former Members of Al-Shabaab," *RUSI - Whitehall Report*. (2019), 4-18.
- Lerner, R.M. et al: "Positive Youth Development, Participation in Community Youth Development Programs, and Community Contributions of Fifth-Grade Adolescents. Findings From the First Wave Of the 4-H Study of Positive Youth Development," *The Journal of Early Adolescence* (2005).
- Möller, K., Küpper, B., Buchheit, F., & Neuscheler, F.: „Evaluation des Aussteigerprogramms für Rechtsextremisten des Landes Nordrhein-Westfalen (APR NRW),“ Esslingen: Ministerium für Inneres und Kommunales (MIK) des Landes Nordrhein-Westfalen (2015).
- Möller, K. & Neuscheler, F.: „Abschlussbericht zur Evaluation der Beratungsstelle Hessen -Religiöse Toleranz statt Extremismus,“ Beratungsstell Hessen Religiöse Toleranz statt Extremismus (2018).
- Phineo: „Kursbuch Wirkung. Das Praxishandbuch für alle, die Gutes noch besser tun wollen,“ 4th edition, 2017.
- Radicalisation Awareness Network: "Ex Post Paper. Guideline Evaluation of PCVE Programmes and Interventions," 2018.

Radicalisation Awareness Network: “Ex Post Paper. Monitoring & Evaluating counter- and alternative narrative campaigns,” 2019a.

Radicalisation Awareness Network: “Evaluating disengagement, deradicalisation and resocialisation efforts,” 2019b.

Radicalisation Awareness Network: “Peer and Self Review Manual for Exit Work,” 2020.

Schuhmacher, Nils [m. Kowol, Greta]: „Evaluation der Fach- und Beratungsstelle für religiös begründete Radikalisierung – Legato. Abschlussbericht“ (PDF). Hamburg: Universität Hamburg, Kriminologische Sozialforschung. (2018).

Uhlmann, M.: “Evaluation of the Advice Centre on Radicalisation. Final Report,” Research Report 31, Nuremberg: Federal Office for Migration and Refugees (2017).

Van der Heide, L. & Schuurmann, B.: “Reintegrating Terrorists in the Netherlands. Evaluating the Dutch approach,” Journal for Deradicalization (2018), No.17, pp.196-238.

Webber, D., Chernikova, M., Kruglanski, A. W., Gelfand, M. J., Hettiarachchi, M., Gunaratna, R., ... & Belanger, J. J.: “Deradicalizing detained terrorists,” Political Psychology (2018), 39(3), pp. 539-556.

ADDITIONAL READING

Christensen, T.W.: “Former Right-Wing Extremists’ Continued Struggle for Self-transformation After an Exit Program. Outlines,” Critical Practice Studies (2019).

Christensen, T.W.: “When trust is essential – civil actors’ role in deradicalization and disengagement initiatives,” in Hansen & Lid (Eds.): Routledge Handbook of Deradicalization and Disengagement (2020), Forthcoming, Manuscript submitted for publication.

Gielen, A.: “Countering Violent Extremism: A Realist Review for Assessing What Works, for Whom, in What Circumstances, and How?,” Terrorism and Political Violence (2019), 31(6), pp. 1149-1167.

Köhler, D.: “Structural Quality Standards for work to intervene with and counter violent extremism. A Handbook for practitioners, state coordination units and civil society programme implementers in Germany,” Competence Centre against Extremism in Baden-Württemberg / Kompetenzzentrum gegen Extremismus in Baden-Württemberg (konex), 2nd edition February 2019.

Kruglanski, A. W., Jasko, K., Chernikova, M., Dugas, M., & Webber, D.: “To the fringe and back. Violent extremism and the psychology of deviance,” American Psychologist (2017), 72(3), pp. 217-230.

Webber, D., Chernikova, M., Molinario, E., & Kruglanski, A. W.: “Psychological approaches to terrorist rehabilitation: Direct and indirect mechanism of deradicalization,” in Hansen & Lid (Eds.): Routledge Handbook of Deradicalization and Disengagement (2020), Forthcoming, Manuscript submitted for publication.



Advancing foreign policy. Since 1955.

Rauchstraße 17/18
10787 Berlin

Tel. +49 30 254231-0

info@dgap.org
www.dgap.org
[@dgapev](#)

The German Council on Foreign Relations (DGAP) is committed to fostering impactful foreign and security policy on a German and European level that promotes democracy, peace, and the rule of law. It is nonpartisan and nonprofit. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the German Council on Foreign Relations (DGAP)

Publisher

Deutsche Gesellschaft für
Auswärtige Politik e.V.

ISSN 1866-9182

Editing Bettina Vestring

Layout & Infografiken

Luise Rombach

Design Concept: WeDo

Author picture(s) © DGAP



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung – Nicht kommerziell – Keine Bearbeitungen 4.0 International Lizenz.